DOCUMENT RESUME

ED 274 705 TM 860 582

AUTHOR van den Berg, Gerald; And Others
TITLE The Usefulness of Comparative Product Research for the Use of Evaluation Results in Curriculum Development.
PUB DATE Apr 86
NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, San Francisco, CA, April 16-20, 1986).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Comparative Analysis; *Curriculum Development; *Curriculum Evaluation; Decision Making; Elementary Secondary Education; Evaluation Methods; *Evaluation Problems; *Evaluation Utilization; Foreign Countries; Hypothesis Testing; *Models; *Predictive Validity; Test Construction
IDENTIFIERS Netherlands

ABSTRACT

This paper examines possible explanations for the fact that evaluation results are rarely used. Most explanatory factors for use of results in decision making mentioned in the literature have been clustered in three categories of variables. They concern: (1) the making of decisions about the practical problem; (2) the way in which evaluation is approached; and (3) the communication between decision makers and evaluation researchers. First, this study reports on a comparative case study in which 14 curriculum development projects were systematically compared with each other. Second, it reports on the experience in three curriculum evaluation projects in which the most important conclusions from the first study were put into practice. The three research projects dealt with curricula for arithmetic/mathematics and Dutch language for elementary education and economics for secondary education in the Netherlands. In each of the research projects, the principal question was whether the difference in efficacy between the curricula was sufficiently marked to justify giving one preference over the other. A competition model was used, the advantages of which were to: (1) force publishers to clarify aims; (2) make evaluation criteria and standards explicit; (3) make research results unconditionally accepted; and (4) confront parties involved with comparisons of their own predictions. It is recommended that the competition model be used in addition to the classical model of empirical testing. (JAZ)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED274705

THE USEFULNESS OF COMPARATIVE PRODUCT RESEARCH FOR THE USE OF EVALUATION RESULTS IN CURRICULUM DEVELOPMENT

Gerard van den Berg
Egbert G. Harskamp
Wim K.B. Hofstee

Institute for Education Research
University of Groningen (RION)
The Netherlands

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
G. van den Berg

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented at the American Research Association Convention, Moscone Center, San Francisco, CA
20 April 1986

BEST COPY AVAILABLE

TM 860 582

Introduction

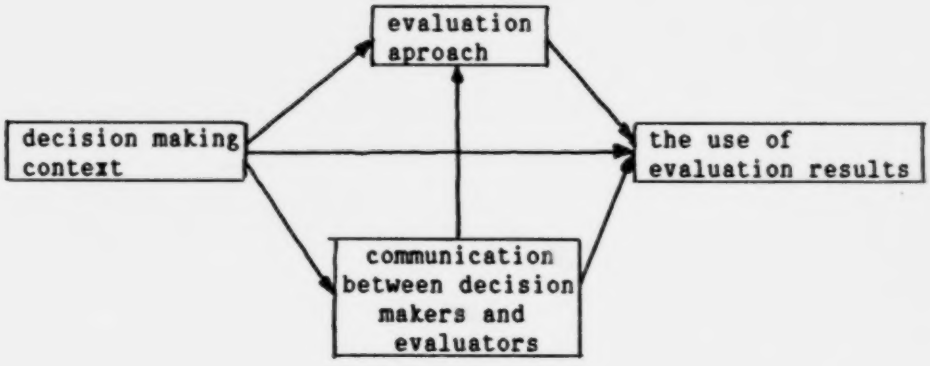
Evaluation research is essentially a type of scientific research concerned with practice, which implies that criteria used to assess to what degree this type of research fulfils its purpose should partly be drawn from the contribution it makes towards solving problems.
This consideration has led many evaluation experts to concern themselves with the probem of the useful application of evaluation results.
It is very difficult to answer the question to what extent there is a real problem of use, because experts work with very divergent definitions of use and because there are no clear standards. Yet the experts have no difference of opinion about the fact that there is a problem of use. But there is no consensus about the question which factors cause the problem of use.
In this paper we shall confine ourselves to the exploration of possible explanations for the fact that evaluation results are rarely used. First we report on a comparative case study in which fourteen curriculum development projects were systematically compared with each other. Secondly we report on the experience in three curriculum-evaluation projects in which the most important conclusions from the first study were put into practice.

Research question

Most explanatory factors for use of results in decisionmaking mentioned in the literature have been clustered in three categories of variables. These constitute the focal concepts in our study. They concern the making of decisions about the practical problem, the way in which evaluation is approached, and the communication between decision makers and evaluation researchers (see: Van den Berg and Hoeben, 1984).

These categories of variables are placed in a conceptional scheme to make it possible to establish their explanatory value (figure 1).

Figure 1: The conceptual scheme of the most important clusters of variables which explain the use of evaluation results.



This conceptual scheme was an important help to define our research questions explicitly:

1. Which definition of use is a valid criterion for judging evaluation research?
2. Which rules of conduct are applied in the various approaches of evaluation research?
3. Which rules of conduct have an impact on the use of evaluation results by decision makers?
4. What is the impact of the decision making context on the communication between the decision makers and evaluators, on the actual application of evaluation rules and on the use of evaluation results by the decision makers?
5. What is the impact of communication between decision makers and evaluators on the actual application of evaluation rules and on the use of evaluation results by the decision makers?
6. Which conclusions can be drawn from the research for improvement of the evaluation practice, so as to produce more useful evaluation results?

The use of evaluation results

An important problem in the discussion about the use of evaluation results in the practice of decision making is the obscurity concerning the most important concept, viz. use.
Patton (1978), Alkin c.s. (1979), Weiss (1979), Leviton Hughes (1981) and others point out this problem.
In the relevant literature about use this concept ranges from a gradual and mainly conceptual influence on decisions, to an immediate and instrumental use of results in decisions.
The choice of the definition of use in the discussion depends on the objective of evaluation research. In our opinion curriculum evaluation research should focus on datacollecting in aid of curriculum improvement. This focus is most valid for formative evaluation research. In this opinion only a restricted definition of use, that is instrumental use, is an acceptable criterion characteristic for good evaluation research.
In this study the amount of use was measured first by content-analysis of several documents, viz. the experimental curriculum (unit), the improved curriculum (unit) and the evaluationreport(s).
Firstly we made an inventory of the modifications in the curriculum (unit). Secondly we explored which modifications were based on or reducible to the evaluation results. Thirdly we set up a standard to identify a fair amount of use. Fourthly we applied this standard to categorize the projects as either frequent or infrequent users of evaluation results in the decision making.
This method was labour-intensive and was done independently by two members of the researchproject.
When there was a difference of opinion we discussed this until consensus was reached. Yet there was a need for control in some way. Therefore we made the decision to ask the opinion of the evaluators and decison makers of the projects about the amount of use in their projects. This complementary method showed only small differences in the judgements of use. Therefore, we may concluded there is a reliable and valid measurement of use in this researchproject.

Approaches to evaluation

Although there is much evaluation that is not research, in the context of this paper evaluation is treated as the equivalent of evaluation research. We distinguish three mainstreams of evaluation approaches. In our attempt to distinguish these three approaches on an operational level we refrain from abstract methodological discussions. Instead we consider evaluation research to be a form of rule-directed conduct, i.e. conduct that is not only according to rules, but which is also the result of the application of rules. In rule-directed conduct the main differences between the two evaluation approaches appear in a concrete and operational way. To that purpose we distinguish rule directed conduct based on the application of technical rules, and rule directed conduct based on the application of comformative rules.
Technical rules are formulated as statements (that can be applied in a concrete way) about how to behave. These rules are instrumental to solving a practical problem. In connection with evaluation research, technical rules fit in with research methods and techniques in the widest sense, such as rules for sampling designs, developing instruments for datacollection and the collecting, processing and analyzing of data. These technics are relatively static, such as mail questionaires and objective tests and scales, closed interviews and statistical techniques.
Conformative rules regulate the comunication between people. These rules demand that each person at least knows the other's meanings. In research this means that the evaluation researchers should know the processes of the attribution of meanings of the research subjects in order to interpret the research results in such a way that research subjects can recognize themselves in them.
On a different level conformative rules concern the sharing of norms and values between people. For evaluation research this means that the evaluation should take into account the norms and values of the person(s) who make the decisions.
Conformative rules are important when collecting information in a more dynamic interaction context. These rules are relatively flexible and dynamic, such as participant observation, open interviews and

qualitative methods for analysis.
It is clear that technical rules in evaluation research are easier to identify than conformative rules. Of course this is closely bound up with the differences in clarity of these rules.
The empirical-analytical approach emphasizes technical rules. In this approach, evaluation research is considered a technical problem. Good communication is necessary, but is is restricted to the formulation of the questions to be answered in the evaluation.
Naturalistic research as well as action research emphasize conformative rules. The naturalistic evaluation approach (see Guba, 1978; Guba and Lincolm, 1981) links up with social-scientific theories which focus on processes of the attribution of meaning in the description of observed reality. It assumes that information always consists of facts interpreted from the unique position and role of the persons researched. It is the researcher's responsibility to interpret this information as accurately as possible. It is an esential part of the conformative system of rules in naturalistic research that the results are presented in such a way, that those who have been researched recognize themselves in the description, and that the decision makers can interpret the description in a meaningful way and base their decision on it.
The action research approach also emphasizes confirmative rules. However, it attaches a different function to these rules.
Communication between decision makers and evaluation researchers is of crucial importance during the entire research process, for the benefit of a continual adjustment of the processes of decisionmaking and research. The decision makers and the evaluators should agree on all decisions that are be made in the course of the research process. To attain this, conformative rules are needed. In that way all the people concerned can participate in a meaningful way when research decisions are at issue. This implies that within evaluation research, technical rules can only be applied if the researchers manage to convey the meaningfulness of these rules to those concerned. Thus, technical rules are less important than conformative rules.
In order tot determine whether evaluators apply either an

empirical-analytical or a naturalistic or action research approach, several project documents, most of them evaluation reports, were analyzed. Complementary information was used from interviews with the evaluators.

Communication between decision makers and evaluators

In our research we concentrate on the initial phase of the evaluation research projects. In this phase it is crucial that decision makers and evaluators should consult each other to answer the question: What information is needed for the solution of decision problems? These consultations begin directly after the decision makers have identified the problem, and after they have determined that the problem cannot, or only with great difficulty, be solved without additional evaluative information. It is often difficult to obtain any clarity concerning the decision makers' precise need for information. This means that the evaluators should identify the decision problem. Nevertheless they often decide for themselves which data they are going to collect, in spite of the fact that the need for information is not specific. A well-considered choice of the data to be collected can only take place on the basis of a clear and concretely formulated decision problem. In this study we concentrate on two output characteristics of the communication process, viz. (1) the nature of the need for information, and (2) the degree to which the questions concerning the evaluation are worked out. Apart from the content of the information needed we should also ask what type of information should be collected, i.e. the question whether there is a need for information that describes and judges, or information that explains and predicts. We submitted this question to the decision makers and evaluators.
The degree to which the questions concerning the evaluation are worked out, is judged on the basis of the various project documents. In a few cases it was impossible to identify the research question(s), because the questions were not formulated. In these cases we labelled the research questions as "not worked-out".
If the research questions were formulated, we judged their degree of concreteness and labelled them either "worked-out" or "not worked-out", depending on the degree of concreteness.

The decision making context

All evaluation research begins with a decision problem. The degree to which decision makers manage to formulate their decisions concretely and clearly is crucial for answering the question about what kind of information is needed to solve the problem. This partly depends on the process of decision making. Within the framework of this study two types of decision making are distinguished - the rational type and the incremental type. In curriculum development all sorts of things should be taken into consideration, e.g. in what manner should the learning process and subject matter of teaching be planned? Curriculum development requires a phased decision making process and consequently drawing up a plan, a framework, for the sake of guiding future decision making. In a rational process of decision making problems are analyzed and divided into the following components: what one wishes to achieve (the goals); what is already available to bridge the gap between the real and the desirable situation (devices); and a description of the context in which some devices are more, and other devices less, effective (see: Simon, 1957). Rational decision making is characterized by taking into account all the facts and arguments that are considered relevant. In an incremental process of decision making one does not analyse the entire decision problem, but only that part of the problem that can be solved in a satisfactory way. After solving this part a new problem is chosen. Decisions are partially made on the basis of coincidences, of trial and eror, of causal learning experiences and pragmatic solutions (see: Lindblom 1968). The two types of decision making mentioned here are theoretical constructs. In reality they do not often occur. However, processes of decision making can be described in terms of 'more or less rational' and 'more or less incremental'.

In order to determine whether curriculum developers approach their subject in a more rational way or in a more incremental way, three indicators are used: (1) the degree to which a problem is diagnosed in a systematic and comprehensive way; (2) the specificity of the decision problem; and (3) the specificity of the need for information.

In order to do so we analyzed various project documents, viz. minutes of project meetings. For a better judgement the findings of these analyses were used in interviews with the projectleaders.
Finally the following guidelines are used for defining the decision making process in terms of 'more rational' or 'more incremental':

1. If the decisionmakers made a systematic and comprehensive diagnosis of the decisionproblems and if the decisionproblem and the need for information was clear and specified, we call the underlying decision making process 'more rational';
2. If the decision makers made an unsystematical and restricted diagnosis of the decision problem and if the decisionproblem and the need for information was loose and unclear, we call the underlying decision making process 'more incremental'.

Results

Although the three evaluation approaches could be discriminated clearly on a conceptual level, it was not easy to identify the projects on empirical levels in terms of the applied rules of conduct in pratice. Only six projects could be identified unambiguously as empirical analytical (a technical approach). The other projects applied predominantly confirmative rules. Three of these projects could not be defined unambiguously in terms of naturalistic or action research. This means that either these evaluators don't know the prevailing rules in these approaches or they disregard the prevailing rules. Consequently we decided to use a dichotomy: a technical or a conformative approach.
Six of the fourteen evaluators stated explicitly that they did not pay attention to the quality of the information (reliability, validity, recognizability and practical intersubjectivity). Three evaluators stated that they did pay attention to the quality of the information but they do not report on this. Only five evaluators pay attention to the quality of the information in their evaluation reports. Evidently a lot of evaluators are under the impression that it is not necessary to report on the quality of the information. The greater part of the

projects neglects to report on the methodical and technical aspects of the evaluation, so that it is extremely difficult to judge the quality of the executed evaluation research. A professional evaluator is more likely to report on the methodical and technical aspects of the evaluation than a non- professional evaluator. In addition, professional evaluators prefer a technical approach. If an evaluator chooses a technical approach the chance is greatest that the evaluation results are used in the decision making process.
It is remarkable that a fair number of projects inside the technical evaluation approach which started the evaluation from a relatively poorly worked-out evaluation problem.
The need for information that explains or predicts is not a sufficient guarantee for starting the evalution from a concretely worked-out evaluation problem. It is not very plausible to assume that the communication between the decision makers and evaluators was focused adequately on the evaluation problem in those projects.
If processes of decision making are of a mainly rational type, the approach in the evaluation study is mostly technical. If processes of decision making are mainly incremental, the approach in the evaluation study is mostly conformative. At the same time it appears that rational decision makers use the results of evaluation research more often than incremental decision makers. The chance of use is greatest when both a mainly rational method of decision making and a technical evaluation approach are applied. On the other hand, the chance of use is slightest when both an incremental method of decision making and a conformative evaluation approach are applied (see appendix: table 1).
If the relation between types of decisionmaking and evaluation approach is checked for the nature of the need of information, specification clearly occurs (see appendix: table 2). Compared with incremental decision making, rational decision making more often needs information that explains or predicts. It is exactly this need that is continually translated into a choice for the more technical evaluation approach. When we attempt to find out how the evaluation approaches influence the use of evaluation results we check this relation successively for the degree to which the evaluation questions are worked out, for the type of information reported, and the quality of

the information reported.
When the questions have been concretely worked out a technical evaluation approach always, and a conformative approach never leads to the use of results. When the questions are not, or hardly, worked out, a conformative approach usually does not lead to the use of results, while a technical approach does so in half of the cases (see appendix: table 3).
When a check was made for the type of information, a specification also occurs. When the evaluation approach is technical there is the greatest chance that not only information is reported that describes and/or judges, but also information that explains and/or predicts; this information is mostly used. When the approach is conformative there is the greatest chance that only information is reported that describes and/or judges, and this information is hardly ever used (see appendix: table 4).
When a check was made for the quality of the information reported, a specification occurs once again (see appendix: table 5). A technical approach always leads to the use of evaluation results, when attention has been paid to demands of quality of the information reported (actually, these were always the demands of reliability and validity). If a technical approach did not pay attention to these demands the effect was that the results were hardly ever used. The same effect occurs when a conformative approach was applied. This was also the case when attention was paid to demands of quality of the information reported (actually, this was always the demand of pratical intersubjectivity).

Consequences for the practice of curriculum evaluation

The most important conclusion is that at least a reasonable degree of use of evaluation results can be predicted if a technical (empirical analytical) approach is used. A technical approach in which the evaluation questions are worked out concretely, criteria of validity and reliability are applied and in which the information supplied also explains and/or predicts instead of just discribes and/or judges (see also Van den Berg and Hoeben, 1984).
To provide a good methodological quality of useful results evaluation

research has to be designed as hypothesis testing research. For, it is a fact that curriculum developers are convinced that the use of their curriculum is the best way to reach a certain goal; and this implies in fact a cause-effect model.

Yet there are little snakes in the grass. Generally curricula pursue many goals and consequently there are many criteria to search for effects. Frequently the evaluator has to make a selection of those goals. Experience shows that many goals cannot, or only with great difficulty, be measured in an objective way. Experience also shows that curriculum developers frequently prefer these goals in evaluation research. In spite of these problems an adequate formulation of criteria is necessary. It is the duty of the curriculum developers to formulate these criteria clearly. Another problem is the choice of evalution standards. In evaluation research it is necessary to judge the effects by previously formulated standards. The absence of standards makes evaluation research impossible. Hofstee (1984 en 1985) argues for the introduction of a-priori-predictions as standards for judging the effectiveness of the curriculum. He takes a further step by making a plan for involving more parties interested in the evaluation research; parties who have a disagreement about the effectiveness of the curriculum.

A difference of opinion is settled in an acceptable manner only when at least one of the parties has been convinced that he ought to budge form his original position. Given that a difference of opinion cannot be settled unilaterally by one party, rules should be developed which are shared by the opposing parties. These rules must enable the parties to decide who is right and who is wrong. For example, if one party makes statements without being able or willing - in the context of acceptability - to appeal to sources recognized by his opponent, and also refuses to defend his statements in any other way, he puts settlement of the difference of opinion beyond the bounds of possibility.

Hofstee (1984 en 1985) pleads for a settlement of differences of opinion by laying bets. This means that the predictions must be verifiable against facts which have yet to be observed. It is important to establish the parties' commitment to the research plan in

order to prevent one of the parties wriggling out of the consequences of the research (Hofstee, 1984). Naturally the option stays open for each party to remain convinced of his own case and to ascribe the research results to chance factors (Dehue, 1984). Right and Wrong lay no claim to truth or untruth. Each party's degree of Rightness or Wrongness depends on the degree to which their predictive assessment of probability (that is the odds laid by them on the possible outcome of the research) meets the actual outcome of the research. The chances of a particular outcome have to be specified in advance because research results are seldom self-evident. Research results then tally better or worse with the advance expectations. Critics state that nobody is willing to stake his reputation without prior knowledge of the evaluation results. In three cases of curriculum evaluation research we have tried to apply Hofstee's recommendations.

The benefit of a comparative product research approach

Although the betting model was the inspiration which led to initation of the three research projects our scenario differed from the betting model on a number of points. In each of the research projects the principal question was whether the difference in efficacy between the curricula was sufficiently marked to justify giving one preference over the other. With such a decision in view, agreement had to be reached with the parties to plan the research as comparative evaluation in which the parties bound themselves beforehand to respect mutually agreed evaluation criteria and standards. This means that:

1. the parties have to reach agreement on the operationalization of both dependent variables;
2. the parties must be asked to estimate the probabilities of possible results; and
3. conditions to decide who is right and who is wrong must be specified beforehand.

In this manner they bind themselves to accept future research results as establishing Rightness of Wrongness.

By setting up the three researches in this way we gave them a certain

competitive character. The characteristic feature of this model is that both parties make predictions of the result of the competition, in absolute as well as comparative terms.

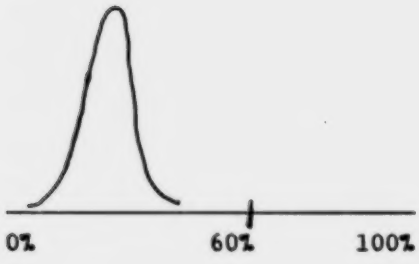
The three research projects dealt with curricula for arithmetic/mathematics and Dutch language for elementary education and economics for secondary education in the Netherlands. The first project included eight educational methods. The second project included five and the third project four published educational methods. All educational methods are published commercially.
From the initial approaches made to the publishers to the time when predictive distributions were formulated, intensive consultation took place. This consultation was always bilateral. Why predictions had to be made required a good deal of explanation. For the sake of testconstructiona lot of consultation was demanded also.
The following procedure was choosen. Firstly we made an inventory of the educational objectives of the methods. Secondly the publishers had to indicate the importance they attached to each objective. On the basis of this information we composed a matrix for (1) stipulating the discrepancies between the methods concerning the pursued objectives, (2) compiling the test. To insure an acceptable psychometric quality of the test we executed several pilot surveys in schools.
Next the publishers were requested to indicate 'how important is this (sub)test for measuring the effectiveness of your method?'. For that purpose a six point ratingscale was used. In the three research projects most publishers judged the (sub)tests for measuring the effectiveness of their methods within the range "fairly important to very important".
In each project the central question to the publishers was 'What percentage of pupils who have been working with your method will get a sufficient score on the sub(test)?'
The publishers were also asked to estimate the range of this percentage. This is the predictive distribution. A pupil passes a (sub)test if he or she answers 55% of the testitems correctly.
For the sake of predicting the percentage of sufficient scores on the (sub)tests the publishers received the overall score of all pupils

from the whole sample (so all methods were included). This score was given for the following reasons:
(1) to push back the chance on gambling and so (2) to get a better check on the claims of the publishers.
It was proposed to ask the publishers whether they thought that other methods would be better, worse of equally good. We felt, however, that this question could not be asked, because of its complexity, so that strictly speaking there was not a true bet laid between the parties (predictions of the difference in quality between their own and competing curricula). But one can point to a less strictly defined bet in the shape of the predictions of results of a single method against the evaluation results actually found in absolute as well as comparative terms.
As we could not simply assume that the methods were being implemented as they were meant to be, we checked whether the minimum demands of implementation were being met. We take implementation to mean 'the application during lesson time by the teachers, of organizational and didactical aspects of the a curriculum, in conformity with the intention of its authors' (see: Van den Berg, 1985).
The organisational aspects were: (1) the amount of lessons given in a regular schoolweek and (2) the degree of differentiation of the curriculumcontent. A didactical aspect was the instruction method (a more interactive method of instruction versus a more frontal instruction method).
We also checked what parts of the method were not or insufficiently studied by the pupils.
For the use of the competition model it is sufficient to check whether or not a method is used by the teachers as intended by the developers. Hofstee (1980) argues that devising all sorts of methodological tests for interfering variables should not automatically be assumed necessary but that, on the contrary, the parties should have good reasons for whatever testing of interfering variables they indulge in. An interfering variable which curriculum evaluation research cannot afford to ignore is the pupils' initial condition. Establishing pupils' initial condition is not simple. Mostly a substitute for initial condition of the pupils must be used e.g. the social

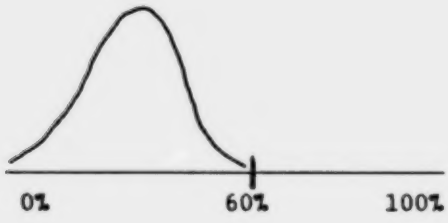
economical status of (one of) the parents.
In order to judge the publishers' predictions, we had to compute a 'degree of right' for each of them. This is done by means of a Bayesian decision rule developed by Hofstee. The consequences of this rule for the publishers are: A publisher who is certain of his prediction can gain much by choosing a narrow predictive distribution if he predicts the correct mean value. If the prediction is wrong however, much is lost. If a publisher chooses a broad predictive distribution than he can neither win nor loose much.
We present an example.

Figure 2: Four hypothetical predictions compared

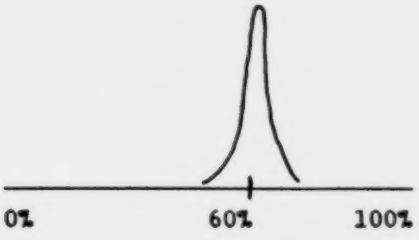
Prediction A



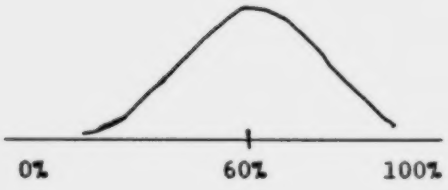
Prediction B



Prediction C



Prediction D



In all these hypothetical cases the score 'empirically' found was 60%. Predictions C en D are accurate. Prediction C, however, has a narrow distribution, so this person who made the prediction was rather certain about the outcome. More certain anyway than predictor D. So prediction C is best, and wins more. The predictions A en B are both wrong, but because of the difference in distribution prediction B is better (less wrong) then prediction A and looses less.
One must realize that a bad prediction on a test does not mean bad testresults, as you can see in situation A. It is possible that the person who made the worst prediction has the best testresults, and the best predictor the worst testresults. This model decides which one has the most realistic claims. Of course the users of the curricula are more interested in the method with the best pupil achievements (testresults). In our research-projects we found rather slight differences in testresults between the methods. Now we need a decision rule that establishes the standards we will accept in order to be able to speak of significant differences.
Therefore we need commitment among publishers about their standards. This was the hardest part in our research projects and we did not succeed in accomplishing this. That is why we had to use the classical hypothesis-testing model. This model appeared to be appropriate for investigating the theoretical expectations which could not be specified in the required a-priori-predictions.

Conclusions

The principal advantages of the competition model are that (1) the publishers are 'forced' to clarify the aims they claim to be pursuing, (2) consequently the evaluation criteria and -standards are implicit, (3) the research results are accepted unconditionally by the parties without resorting to a whole complex of immunization strategies, (4) the parties will be more inclined to reflect because they are confronted with a comparison of their own predictions with the results of research. It is recommendable to use the competition model in addition to the classical model of empirical testing.

REFERENCES

Alkin, M.C., Daillak, R. en P. White (1979). Using evaluation: Does evaluation make a difference? Beverly Hills: Safe.

Berg, G. van den (1985). Curriculum evaluation as comparative product research. In: Creemers, B.P.M. (ed). Evaluation research in education, p.23-35. The Hague: SVO.

Berg, G. van den en W.Th.J.G. Hoeben (1984). The practice of evaluation research and the use of evaluation results. In: Studies in Educational Evaluation. Vol. 10, p.309-321.

Dehue, T. (1984). Rationaliteit bij incommensurabiliteit. In: Kennis en Methode, p.17-24.

Guba, E.G. (1978). Toward a methodology of naturalistic inquiry in educational evaluation. Los Angeles.

Guba, E.G. en Y.S. Lincoln (1981). Effective evaluation. Los Angeles.

Hofstee, W.K.B. (1984). Methodological decision rules as research. policies: a betting model reconstruction of empirical research. In: Acta Psychologica, International Journal of Psychonomics, n. .56, p.93-109.

Hofstee, W.K.B. (1985). A betting model of evaluation research. In Creemers, B.P.M. (ed). Evaluation Research in Education, p.15-21. The Hague: SVO.

Leviton, L.C. en E.F.K. Hughes (1981). Research on the utilization of evaluation. In: Educational Research, no 5, p.525-548.

Lindblom, C.E. (1985). The intelligence of democracy: decision making through mutual adjustment. Free Press.

Patton, M.Q. (1978). Utilization- forused research. Beverly Hills: Safe.

Simon, M.A. (1957). Administrative Behaviour. Macoudlan.

Weiss, G.H. (1979). The many meanings of research utilization. In: Public Administration Review, no. 5.p.426-431.